Ranking Individuals by Group Comparisons

Chih-Jen Lin
Department of Computer Science
National Taiwan University

Joint work with Tzu-Kuo Huang and Ruby C. Weng

Outline

- Problem

  Ranking individuals by group comparisons
- Existing approaches
- New approaches
- Real applications:

  Ranking bridge partnerships

  Multi-class classification by error-correcting codes
- Discussion and conclusions

The Problem

- Many sports are team comparisons
  How to rank individuals?
- Rank a basketball player by average points
  But ignore teammates'/opponents' abilities.
- In bridge
  Two partnerships vs. two partnerships
  Match record shows which two are better
  But how to rank partnerships?

- Multi-class classification by error-correcting codes (Dietterich and Bakiri 1995; Allwein et al. 2001)

  Some classes vs. some others

  Finding the winning class (individual)

Existing Work: Huang et al. (NIPS 2004)

- Games:
  $\{1,3\}$ vs. $\{2\}$; $\{1,4\}$ vs. $\{2,3,5\}$, etc.
- Individual j's ability: $p_j \geq 0$
- The ith setting: team I_i^+ vs. team I_i^-

$$P(I_i^+ \text{ beats } I_i^-) = \frac{\sum_{j:j \in I_i^+} p_j}{\sum_{j:j \in I_i} p_j}.$$

Existing Work: Huang et al. (NIPS 2004)

- Games:
  $\{1,3\}$ vs. $\{2\}$; $\{1,4\}$ vs. $\{2,3,5\}$, etc.
- Individual j's ability: $p_j \geq 0$
- The ith setting: team I_i^+ vs. team I_i^-

$$P(I_i^+ \text{ beats } I_i^-) = \frac{\sum_{j:j\in I_i^+} p_j}{\sum_{j:j\in I_i} p_j}.$$

- Extension of Bradley-Terry model (1952):

$$P(\text{individual } i \text{ beats individual } j) = \frac{p_i}{p_i + p_j}$$

- Minimizing the negative log-likelihood

$$\min_{\mathbf{p}} \quad -\sum_{i=1}^{m} \left(n_i^+ \log \frac{\sum_{j:j\in I_i^+} p_j}{\sum_{j:j\in I_i} p_j} + n_i^- \log \frac{\sum_{j:j\in I_i^-} p_j}{\sum_{j:j\in I_i} p_j} \right)$$

$$\text{subject to} \quad \sum_{j=1}^{k} p_j = 1, 0 \le p_j, j = 1, \ldots, k.$$

- n_i^+ and n_i^-: # wins by I_i^+ and I_i^-; $\quad n_i \equiv n_i^+ + n_i^-$
- May be non-convex
- Used for multi-class prob. outputs in our SVM software

A Naive Approach: SUM

- Summing # of winning games

$$\frac{\sum_{i:s\in I_i^+} n_i^+ + \sum_{i:s\in I_i^-} n_i^-}{\sum_{i:s\in I_i} 1}.$$

- Not consider opponents' abilities

  Susceptible to individuals playing very few (or many) games
- Not consider teammates' abilities

  Strong and weak players: the same credits
- Ranking by SUM similar to that of teams.

A New Exponential Model

- k individuals' abilities: a vector $\mathbf{v} \in R^k$, $-\infty < v_s < \infty$, $s = 1, \ldots, k$.
- Ability of team I_i^+ and I_i^-: sum of members'

$$T_i^+ \equiv \sum_{s:s \in I_i^+} v_s \quad \text{and} \quad T_i^- \equiv \sum_{s:s \in I_i^-} v_s.$$

- Teams' actual performances: random variables Y_i^+ and Y_i^-

$$P(I_i^+ \text{ beats } I_i^-) \equiv P(Y_i^+ - Y_i^- > 0).$$

- Distribution unknown, assume doubly-exponential extreme-value distribution

$$P(Y_i^+ \leq y) = \exp(-e^{-(y - T_i^+)}),$$

- Probability that I_i^+ wins

$$P(I_i^+ \text{ beats } I_i^-) = \frac{e^{T_i^+}}{e^{T_i^+} + e^{T_i^-}}.$$

- Can assume normal; but model more complex

Estimation: Regularized Least Square

- n_i^+ and n_i^-: # games teams I_i^+ and I_i^- win

$$\frac{e^{T_i^+}}{e^{T_i^+} + e^{T_i^-}} \approx \frac{n_i^+}{n_i^+ + n_i^-} \quad \Rightarrow \quad e^{T_i^+ - T_i^-} \approx \frac{n_i^+}{n_i^-}.$$

- Solve

$$\min_{\mathbf{v}} \quad \sum_{i=1}^{m} ((T_i^+ - T_i^-) - \log(n_i^+/n_i^-))^2$$

- Unique solution:

Adding a regularized term $\mu \mathbf{v}^T \mathbf{v}$

Estimation: Maximum Likelihood (ML)

- Negative log-likelihood function:

$$\arg\min l(\mathbf{v})$$

where

$$l(\mathbf{v}) \equiv -\sum_{i=1}^{m} \left(n_i^+ \log \frac{e^{T_i^+}}{e^{T_i^+} + e^{T_i^-}} + n_i^- \log \frac{e^{T_i^-}}{e^{T_i^+} + e^{T_i^-}} \right)$$

- Convex

Maximum Likelihood: Iterative Scaling

- Standard optimization methods: Newton's etc.

Maximum Likelihood: Iterative Scaling

- Standard optimization methods: Newton's etc.
- Simple iterative scaling

  Update the sth component;

  $\boldsymbol{\delta} \equiv [0, \ldots, 0, \delta_s, 0, \ldots, 0]^T$

$$l(\mathbf{v} + \boldsymbol{\delta}) - l(\mathbf{v}) \leq -\left(\sum_{i:s \in I_i^+} n_i^+ + \sum_{i:s \in I_i^-} n_i^- \right) \delta_s + \left(\sum_{i:s \in I_i^+} \frac{n_i e^{T_i^+}}{e^{T_i^+} + e^{T_i^-}} + \sum_{i:s \in I_i^-} \frac{n_i e^{T_i^-}}{e^{T_i^+} + e^{T_i^-}} \right) (e^{\delta_s} - 1)$$

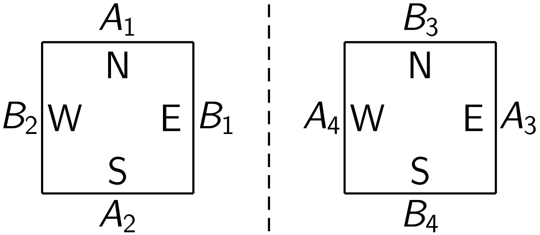
- Updating rule

$$v_s \leftarrow v_s + \log \frac{\sum_{i:s\in I_i^+} n_i^+ + \sum_{i:s\in I_i^-} n_i^-}{\sum_{i:s\in I_i^+} \frac{n_i e^{T_i^+}}{e^{T_i^+}+e^{T_i^-}} + \sum_{i:s\in I_i^-} \frac{n_i e^{T_i^-}}{e^{T_i^+}+e^{T_i^-}}}.$$

- Under a minor assumption,
  any limit point a global minimum

Ranking Bridge Partnerships

- A match setting:



Team A: two partnerships (A_1, A_2), (A_3, A_4)

Team B: two partnerships (B_1, B_2), (B_3, B_4)

- Same board for (B_1, B_2), (A_3, A_4)

Avoid uneven hands

Bridge Scoring

- First three boards; India vs. Portugal in Bermuda Bowl 2005

Board	Table I		Table II		IMPs	
	NS	EW	NS	EW	IN	PT
1		1510		1510		
2	100		650			11
3		630		630		

- International Match Points (IMPs): difference in two teams' total scores
- IMPs of all rounds ⇒ Victory Points (VP)

  Overall results between two teams; used as n_i^+, n_i^-

Experimental Settings

- Bermuda Bowl 2005, the most prestigious bridge event
- 22 teams, round robin
  $\binom{22}{2} = 231$ matches
- Most teams: six players $\Rightarrow$ three fixed partnerships playing similar # matches
- Total 69 partnerships
  Other details not shown here

Results

Team	Partnership rankings											
	RLS			ML			HNG			SUM		
IT	21	18	13	8	17	18	7	14	16	5	4	12
US2	63	67	1	52	66	1	43	66	1	47	29	2
US1	9	36	41	10	19	37	10	15	38	23	6	10
SE	2	55	37	2	25	53	2	12	64	1	19	39
IN	14	40	42	9	32	41	9	30	42	20	14	15
AR	33	26	32	26	21	29	25	23	34	16	17	28
EG	47	30	27	43	27	13	52	22	13	38	22	3
	46			57			50			8		
⋯												
AU	44	51	13	42	51	20	40	53	21	42	51	41
NZ	68	23	3	68	48	5	67	39	4	64	36	13
UK	10	24	59	12	36	64	17	33	63	45	18	54
⋯												

Analysis: Are Obtained Rankings Good?

- Two teams (r_1, r_2) and $(\bar{r}_1, \bar{r}_2)$. Define

$$r \text{ better than } \bar{r} \text{ if } \max(r_1, r_2) < \min(\bar{r}_1, \bar{r}_2).$$

- Example: first two partnerships of Italy and USA2

  (21, 18) better than (63, 67)

Analysis: Are Obtained Rankings Good?

- Two teams (r_1, r_2) and $(\bar{r}_1, \bar{r}_2)$. Define

$$r \text{ better than } \bar{r} \text{ if } \max(r_1, r_2) < \min(\bar{r}_1, \bar{r}_2).$$

- Example: first two partnerships of Italy and USA2

  (21, 18) better than (63, 67)

- Violation: r better than $\bar{r}$ but $\bar{r}$ beats r
- Number of violations

RLS	ML	HNG	SUM
12	6	9	32

- Other details not shown

Top 10 partnerships by the approach ML

Team	Players	
U.S.A.2	Eric Greco	Geoff Hampson
Sweden	Peter Bertheau	Fredrik Nystrom
Japan	Yoshiyuki Nakamura	Yasuhiro Shimizu
Chinese Taipei	Chih-Kuo Shen	Jui-Yiu Shih
New Zealand	Tom Jacob	Malcolm Mayer
China	Zhong Fu	Jie Zhao
Brazil	Gabriel Chagas	Miguel Villas-boas
Italy	Norberto Bocchi	Giorgio Duboin
India	Subhash Gupta	Rajeshwar Tewari
U.S.A.1	Jeff Meckstroth	Eric Rodwell

Multi-class Classification via ECOC

- The same settings as Huang et al. (NIPS 2004)
- Error rates: new models are competitive

		Dense			Sparse		
Problem	#class	RLS	ML	HNG	RLS	ML	HNG
dna	3	6.35	6.34	6.39	6.88	6.29	6.24
waveform	3	13.71	13.71	13.92	13.54	13.45	14.27
satimage	6	11.61	11.52	11.41	11.46	11.58	11.79
segment	7	3.54	3.46	3.45	3.97	3.54	3.23
USPS	10	7.22	7.29	7.66	8.06	7.68	8.52
MNIST	10	7.25	7.25	7.58	8.09	7.74	8.97
letter	26	19.55	19.37	20.27	21.20	20.47	20.43

Conclusions

- New and useful methods to rank individuals from group comparisons

  Convex formulations
- Experiments:

  Bridge partnership rankings

  Multi-class classification by error-correcting codes
- Techniques to evaluate different rankings